(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 October 2000 (19.10.2000)

PCT

(10) International Publication Number
WO 00/62502 A3

(51) International Patent Classification[7]: H04L 29/06

(21) International Application Number: PCT/US00/09861

(22) International Filing Date: 12 April 2000 (12.04.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/128,872	12 April 1999 (12.04.1999)	US
09/437,637	10 November 1999 (10.11.1999)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	09/437,637 (CIP)
Filed on	10 November 1999 (10.11.1999)
US	60/128,872 (CIP)
Filed on	12 April 1999 (12.04.1999)

(71) Applicant *(for all designated States except US)*: **RAINFINITY, INC.** [US/US]; Suite 200, 87 N. Raymond Avenue, Pasadena, CA 91103 (US).

(72) Inventors; and

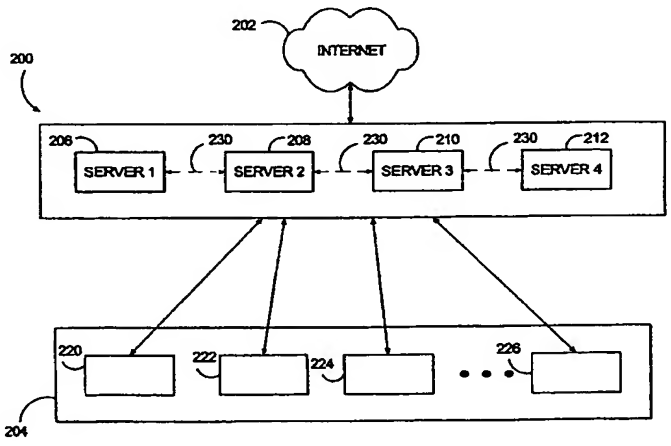
(75) Inventors/Applicants *(for US only)*: **BRUCK, Jehoshua** [US/US]; 5657 Bramblewood Road, La Canada, CA 91011 (US). **BOHOSSIAN, Vasken** [CA/US]; 1127 E. Del Mar Boulevard #227, Pasadena, CA 91106 (US). **FAN, Chenggong** [CN/US]; 1155 E. Del Mar Boulevard #105, Pasadena, CA 91106 (US). **LEMAHIEU, Paul** [US/US]; 1032 E. Del Mar Boulevard #301, Pasadena, CA 91106 (US). **LOVE, Philip** [GB/US]; 1032 E. Del Mar Boulevard #301, Pasadena, CA 91106 (US).

(74) Agents: **HALL, David, A.** et al.; Heller Ehrman White & McAuliffe LLP, Suite 700, 4250 Executive Square, La Jolla, CA 92037 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: DISTRIBUTED SERVER CLUSTER FOR CONTROLLING NETWORK TRAFFIC



(57) Abstract: A scalable, distributed, highly available, load balancing server system having multiple machines is provided that functions as a front server layer between a network (such as the Internet) and a back-end server layer having multiple machines functioning as Web file servers, FTP servers, or other application servers. The front layer machines comprise a server cluster that performs fail-over and dynamic load balancing for both server layers. The operation of the servers on both layers is monitored, and when a server failure at either layer is detected, the system automatically shifts network traffic from the failed machine to one or more operational machines, reconfiguring front-layer servers as needed without interrupting operation of the server system. The server system automatically accommodates additional machines in the server cluster, without service interruption. The system operates with a dynamic reconfiguration protocol that permits reassignment of network addresses to the front layer machines. The front layer machines perform their operations without breaking network communications between clients and servers, and without rebooting of computers.

WO 00/62502 A3

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *With international search report.*

(88) **Date of publication of the international search report:**
29 March 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/US 00/09861

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04L29/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, IBM-TDB, INSPEC, COMPENDEX

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 26559 A (GTE INTERNETWORKING INC) 18 June 1998 (1998-06-18) page 13, line 29 -page 21, line 23 figures 2B,3C,4C	1,2
A	---	3,18,30
A	US 5 341 477 A (MORENCY JOHN P ET AL) 23 August 1994 (1994-08-23) figures 2,2A,7 column 4, line 38 - line 68 column 11, line 15 - line 50	1-3,18, 30
A	--- US 5 774 668 A (CHOQUIER PHILIPPE ET AL) 30 June 1998 (1998-06-30) figures 1,5A,5B column 4, line 54 -column 8, line 63 column 12, line 5 -column 13, line 38	1-3,18, 30
	--- -/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 September 2000	06/10/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Eraso Helguera, J

1

Form PCT/ISA/210 (second sheet) (July 1992)

International Application No
PCT/US 00/09861

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages.	Relevant to claim No.
A	US 5 774 660 A (LIU ZAIDE ET AL) 30 June 1998 (1998-06-30) figure 19 column 18, line 43 -column 19, line 45	1-30
A	GOLDSZMIDT G S: "LOAD MANAGEMENT FOR SCALING UP INTERNET SERVICES" IEEE NETWORK OPERATIONS AND MANAGEMENT SYMPOSIUM,US,NEW YORK, NY: IEEE, vol. CONF. 10, 15 February 1998 (1998-02-15), pages 828-835, XP000793430 ISBN: 0-7803-4352-2 the whole document	1-30
P,X	WO 99 33227 A (HOLONTECH CORP) 1 July 1999 (1999-07-01) page 11, paragraph 2 -page 12, paragraph 3 page 16, paragraph 2 -page 18, paragraph 2 figure 2	1
P,X	US 5 898 830 A (COLEY CHRISTOPHER D ET AL) 27 April 1999 (1999-04-27) figures 4,9 column 8, line 25 -column 9, line 51 column 12, line 29 -column 13, line 53	1,2

1

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9826559 A	18-06-1998	AU 5692498 A	03-07-1998
		EP 1016253 A	05-07-2000
US 5341477 A	23-08-1994	AT 151183 T	15-04-1997
		AU 611605 B	13-06-1991
		AU 4996190 A	13-09-1990
		AU 630291 B	22-10-1992
		AU 7603391 A	15-08-1991
		CA 2010762 A	24-08-1990
		DE 69030340 D	07-05-1997
		DE 69030340 T	20-11-1997
		EP 0384339 A	29-08-1990
		JP 3116262 A	17-05-1991
US 5774668 A	30-06-1998	US 5951694 A	14-09-1999
US 5774660 A	30-06-1998	NONE	
WO 9933227 A	01-07-1999	AU 1803099 A	12-07-1999
US 5898830 A	27-04-1999	US 6052788 A	18-04-2000

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : H04L 29/00 — A2

(11) International Publication Number: **WO 00/62502**

(43) International Publication Date: 19 October 2000 (19.10.00)

(21) International Application Number: PCT/US00/09861

(22) International Filing Date: 12 April 2000 (12.04.00)

(30) Priority Data:

60/128,872	12 April 1999 (12.04.99)	US
09/437,637	10 November 1999 (10.11.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	09/437,637 (CIP)
Filed on	10 November 1999 (10.11.99)
US	60/128,872 (CIP)
Filed on	12 April 1999 (12.04.99)

(71) Applicant *(for all designated States except US)*: RAINFINITY, INC. [US/US]; Suite 200, 87 N. Raymond Avenue, Pasadena, CA 91103 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: BRUCK, Jehoshua [US/US]; 5657 Bramblewood Road, La Canada, CA 91011 (US). BOHOSSIAN, Vasken [CA/US]; 1127 E. Del Mar Boulevard #227, Pasadena, CA 91106 (US). FAN, Chenggong [CN/US]; 1155 E. Del Mar Boulevard #105, Pasadena, CA 91106 (US). LEMAHIEU, Paul [US/US]; 1032 E. Del Mar Boulevard #301, Pasadena, CA 91106 (US). LOVE, Philip [GB/US]; 1032 E. Del Mar Boulevard #301, Pasadena, CA 91106 (US).

(74) Agents: HALL, David, A. et al.; Heller Ehrman White & McAuliffe LLP, Suite 700, 4250 Executive Square, La Jolla, CA 92037 (US).

(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: DISTRIBUTED SERVER CLUSTER FOR CONTROLLING NETWORK TRAFFIC

(57) Abstract

A scalable, distributed, highly available, load balancing server system having multiple machines is provided that functions as a front server layer between a network (such as the Internet) and a back-end server layer having multiple machines functioning as Web file servers, FTP servers, or other application servers. The front layer machines comprise a server cluster that performs fail-over and dynamic load balancing for both server layers. The operation of the servers on both layers is monitored, and when a server failure at either layer is detected, the system automatically shifts network traffic from the failed machine to one or more operational machines, reconfiguring front-layer servers as needed without interrupting operation of the server system. The server system automatically accommodates additional machines in the server cluster, without service interruption. The system operates with a dynamic reconfiguration protocol that permits reassignment of network addresses to the front layer machines. The front layer machines perform their operations without breaking network communications between clients and servers, and without rebooting of computers.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

CLAIMS

We claim:

1. A data traffic controller for a computer network, the controller comprising:

a network interface that permits communication between the traffic controller and a subnet over which network data is sent and received; and

a distributed gateway application that dynamically determines network data traffic address assignments from multiple primary network addresses to multiple virtual network addresses to send network data to an intended host on the subnet, wherein network data intended for a host on the subnet is addressed to one of the virtual network addresses.

2. A method of controlling data traffic for a computer network through a traffic controller computer, wherein the data traffic is received through a network interface that permits communication between the traffic controller and a subnet over which network data is sent and received, the method comprising:

receiving network data intended for a host on the subnet at a data traffic controller, wherein the network data is addressed to one of a plurality of virtual network addresses that are on the subnet and are associated with one or more primary network addresses; and

dynamically determining network data traffic address assignments from multiple primary network addresses to multiple virtual network addresses to send network data to an intended host on the subnet.

3. A method of operating a server computer for controlling data traffic of a computer network, the method comprising:

receiving network data traffic through a network interface that permits communication between the server computer and other computers;

communicating with a plurality of server computers that are all members of a first subnet of network addresses over which network data is sent and received, comprising a front layer of servers, wherein the communication includes state sharing information with a dynamic reconfiguration protocol that permits reassignment of network addresses among the front layer servers and specifies state information sharing and load information sharing among the front layer servers; and

communicating with a plurality of network computers that are members of a second subnet of network addresses to send and receive network data traffic.

4. A method as defined in claim 3, wherein communicating with a plurality of server computers comprises sending data using a Reliable Message layer scheme that comprises a token data packet and one or more data carriage packets, wherein the token data packet specifies the number of data carriage packets that together comprise a Reliable

Message packet and wherein the data carriage packets include data relating to state information and data traffic load information about each of the front layer servers.

5. A method as defined in claim 3, further comprising:

receiving network data traffic;

determining if the data traffic is associated with a previous network communication session of an original server computer of the first subnet, prior to a network address reassignment for the original server computer;

responding to data traffic not associated with a previous network communications session of an original server computer by processing the data traffic; and

responding to data traffic that is associated with a previous network communication session with an original server computer by checking a record of network address assignments and identifying the original server computer associated with the previous network communications session and forwarding the data traffic to the identified original server computer.

6. A method as defined in claim 3, wherein communicating with server computers of the first subnet further includes performing a network address translation comprising:

receiving data traffic for a pool of virtual network addresses serviced by the server computers of the first subnet;

determining that the received data traffic includes a data packet intended for a port connection at a different server computer of the first subnet; and

identifying a computer port assignment of the different server computer in the first subnet for which the data traffic is intended and performing an address translation function to route the data packet to the different server computer.

7. A method as defined in claim 6, wherein determining a port connection of the received data traffic comprises determining that the data packet relates to a port connection that is not in a list of port connections, and wherein identifying a port assignment comprises receiving a synchronization message update containing port assignment information that permits identification of the different server computer to which the port is assigned.

8. A method as defined in claim 3, further including:

receiving cluster configuration information for operation of the server computer and adapting operation accordingly; and

communicating the cluster configuration information to the other server computers of the first subnet such that the other server computers adapt their operation accordingly.

9. A method as defined in claim 3, further comprising:

receiving data traffic comprising a request for a data file;

sending a data packet with the request information to a computer of the second subnet;

storing header information for the data request;

receiving data packets of the requested data file from the second subnet computer and forwarding the data packets to the requesting computer;

maintaining state data on the client communications session, including the number of data packets sent to the requesting computer;

detecting a failure of the second subnet computer and in response identifying a replacement second subnet computer from which the requested data is available; and

sending a request for the requested data to the replacement second subnet computer, such that the request is for data beginning subsequent to the data packets already forwarded to the requesting computer.

10. A method as defined in claim 3, further including:

configuring an operating system of the server computer such that all network addresses in a pool of addresses assigned to the server computers of the first subnet are assigned to the server computer;

generating a gratuitous address resolution protocol (ARP) message in response to an address reassignment of the server computer and communicating the ARP message to the other server computers of the first subnet;

blocking the sending of an ARP acknowledgment message to the other server computers of the first subnet for any received gratuitous ARP message, thereby inhibiting reboot operation of the respective server computers and ensuring that each server computer is unaware of any duplicate assignment of network address numbers.

11. A method as defined in claim 3, further including operating as an authoritative node of the first subnet to ensure symmetric routing of network data traffic to and from the first subnet.

12. A method as defined in claim 11, wherein operating to ensure symmetric traffic routing comprises:

receiving a data request from a responding server computer of the first subnet, wherein the data request was initially received at the responding server computer, which determined the authoritative node for responding to the data request;

identifying a server computer in the first subnet that will handle the data traffic associated with the data request and forwarding the data request to the identified server computer for handling;

receiving a reply message from a server computer of the first subnet that is operating as a default reply node to a second subnet computer that is responding to the data request; and

forwarding the reply message to a server computer of the first subnet that will ensure symmetric routing of the data request and reply message with respect to the server computers of the first subnet.

13. A method as defined in claim 12, further comprising forwarding assignment information to the server computer of the first subnet that was operating as the default reply node for the data request, wherein the assignment information includes forwarding information that the default reply node can use to directly forward response messages from the second subnet computer to the first subnet computer that will ensure symmetric routing.

14. A method as defined in claim 3, wherein the computers of the second subnet comprise application servers.

15. A method as defined in claim 3, wherein the network over which data traffic is received comprises the Internet.

16. A method as defined in claim 15, wherein the network data traffic includes requests for data files.

17. A method as defined in claim 16, wherein the data files comprise Web pages.

18. A program product for use in a computer that executes program steps recorded in a computer-readable media to perform a method of operating the computer for controlling data traffic of a computer network, the program product comprising:

a recordable media;

computer-readable instructions recorded on the recordable media, comprising instructions executable by the computer to perform a method comprising:

receiving network data traffic through a network interface that permits communication between the server computer and other computers;

communicating with a plurality of server computers that are all members of a first subnet of network addresses over which network data is sent and received, comprising a front layer of servers, wherein the communication includes state sharing information with a dynamic reconfiguration protocol that permits reassignment of network addresses among the front layer servers and specifies state information sharing and load information sharing among the front layer servers; and

communicating with a plurality of network computers that are members of a second subnet of network addresses to send and receive network data traffic.

19. A program product as defined in claim 18, wherein communicating with a plurality of server computers comprises sending data using a Reliable Message layer scheme that comprises a token data packet and one or more data carriage packets, wherein the token data packet specifies the number of data carriage packets that together comprise a Reliable

Message packet and wherein the data carriage packets include data relating to state information and data traffic load information about each of the front layer servers.

20. A program product as defined in claim 18, wherein the performed method further comprises:

receiving network data traffic;

determining if the data traffic is associated with a previous network communication session of an original server computer of the first subnet, prior to a network address reassignment for the original server computer;

responding to data traffic not associated with a previous network communications session of an original server computer by processing the data traffic; and

responding to data traffic that is associated with a previous network communication session with an original server computer by checking a record of network address assignments and identifying the original server computer associated with the previous network communications session and forwarding the data traffic to the identified original server computer.

21. A program product as defined in claim 18, wherein communicating with server computers of the first subnet further includes performing a network address translation comprising:

receiving data traffic for a pool of virtual network addresses serviced by the server computers of the first subnet;

determining that the received data traffic includes a data packet intended for a port connection at a different server computer of the first subnet; and

identifying a computer port assignment of the different server computer in the first subnet for which the data traffic is intended and performing an address translation function to route the data packet to the different server computer.

22. A program product as defined in claim 21, wherein determining a port connection of the received data traffic comprises determining that the data packet relates to a port connection that is not in a list of port connections, and wherein identifying a port assignment comprises receiving a synchronization message update containing port assignment information that permits identification of the different server computer to which the port is assigned.

23. A program product as defined in claim 18, wherein the performed method further includes:

receiving cluster configuration information for operation of the server computer and adapting operation accordingly; and

communicating the cluster configuration information to the other server computers of the first subnet such that the other server computers adapt their operation accordingly.

24. A program product as defined in claim 18, wherein the performed method further comprises:

receiving data traffic comprising a request for a data file;

sending a data packet with the request information to a computer of the second subnet;

storing header information for the data request;

receiving data packets of the requested data file from the second subnet computer and forwarding the data packets to the requesting computer;

maintaining state data on the client communications session, including the number of data packets sent to the requesting computer;

detecting a failure of the second subnet computer and in response identifying a replacement second subnet computer from which the requested data is available; and

sending a request for the requested data to the replacement second subnet computer, such that the request is for data beginning subsequent to the data packets already forwarded to the requesting computer.

25. A program product as defined in claim 18, wherein the performed method further includes:

configuring an operating system of the server computer such that all network addresses in a pool of addresses assigned to the server computers of the first subnet are assigned to the server computer;

generating a gratuitous address resolution protocol (ARP) message in response to an address reassignment of the server computer and communicating the ARP message to the other server computers of the first subnet;

blocking the sending of an ARP acknowledgment message to the other server computers of the first subnet for any received gratuitous ARP message, thereby inhibiting reboot operation of the respective server computers and ensuring that each server computer is unaware of any duplicate assignment of network address numbers.

26. A program product as defined in claim 18, wherein the performed method further includes operating as an authoritative node of the first subnet to ensure symmetric routing of network data traffic to and from the first subnet.

27. A program product as defined in claim 26, wherein operating to ensure symmetric traffic routing comprises:

receiving a data request from a responding server computer of the first subnet, wherein the data request was initially received at the responding server computer, which determined the authoritative node for responding to the data request;

identifying a server computer in the first subnet that will handle the data traffic associated with the data request and forwarding the data request to the identified server computer for handling;

receiving a reply message from a server computer of the first subnet that is operating as a default reply node to a second subnet computer that is responding to the data request; and

forwarding the reply message to a server computer of the first subnet that will ensure symmetric routing of the data request and reply message with respect to the server computers of the first subnet.

28. A program product as defined in claim 27, wherein the performed method further comprises forwarding assignment information to the server computer of the first subnet that was operating as the default reply node for the data request, wherein the assignment information includes forwarding information that the default reply node can use to directly forward response messages from the second subnet computer to the first subnet computer that will ensure symmetric routing.

29. A program product as defined in claim 18, wherein the computers of the second subnet comprise application servers.

30. A network server computer comprising:

a network interface that permits communication between the server computer and other computers;

a distributed server application executed by the server computer that thereby permits the server computer to communicate with a plurality of server computers that are all members of a first subnet of network addresses over which network data is sent and received, comprising a front layer of servers, wherein the communication includes state sharing information with a dynamic reconfiguration protocol that permits reassignment of network addresses among the front layer servers and specifies state information sharing and load information sharing among the front layer servers, and permits the server computer to communicate with a plurality of network computers that are members of a second subnet of network addresses to send and receive network data traffic.

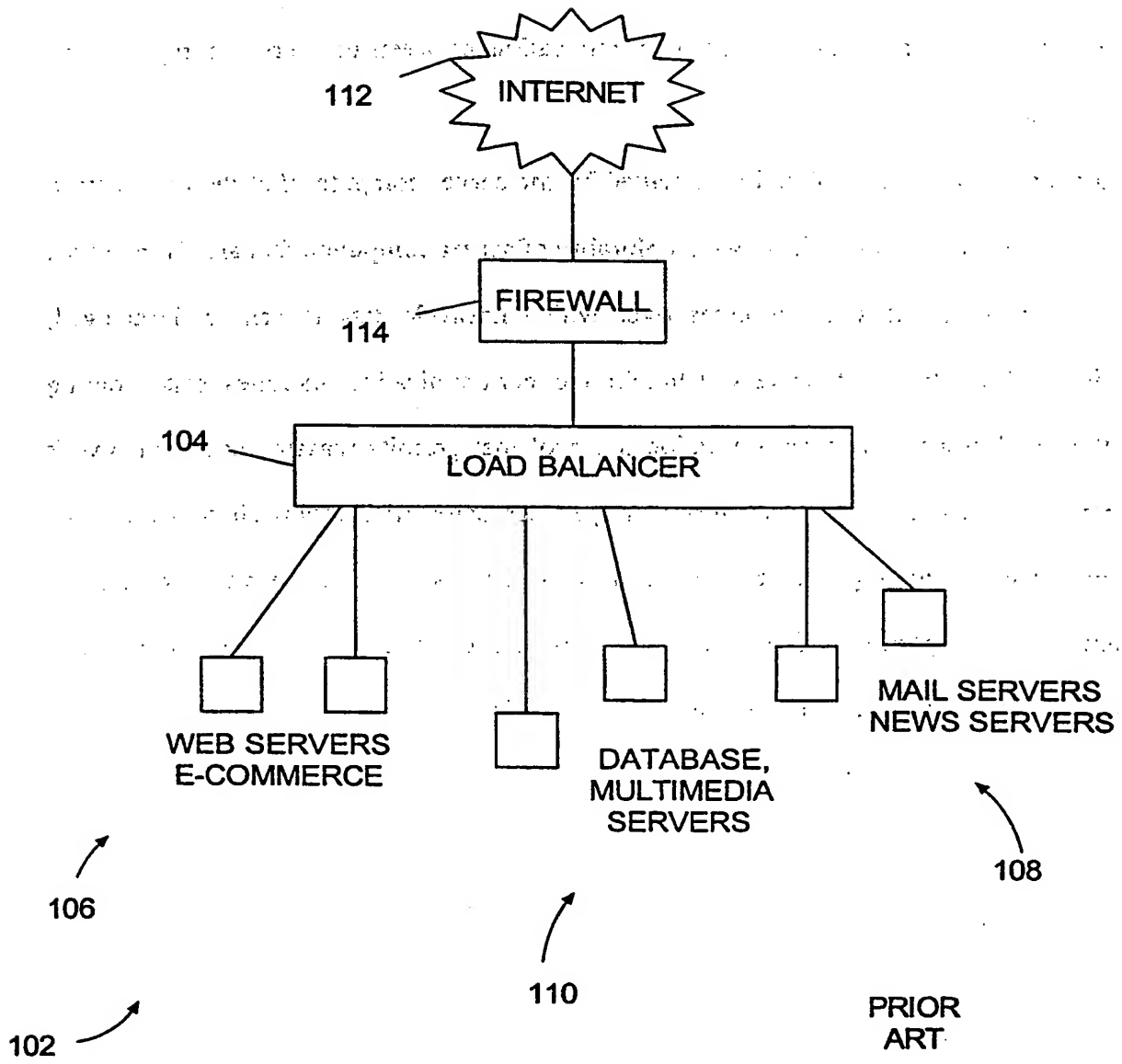


FIG. 1

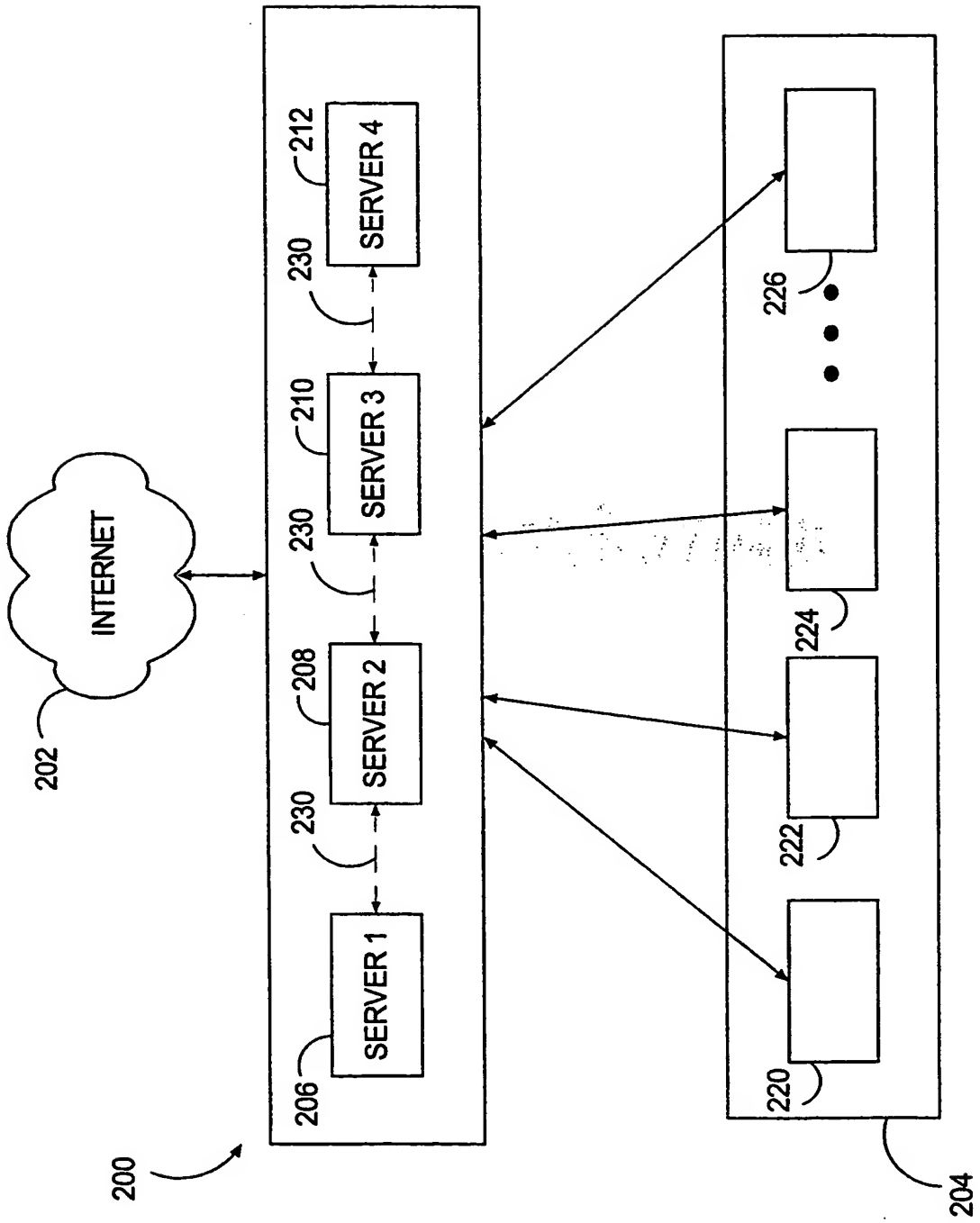


FIG. 2

This Page Blank (uspto)

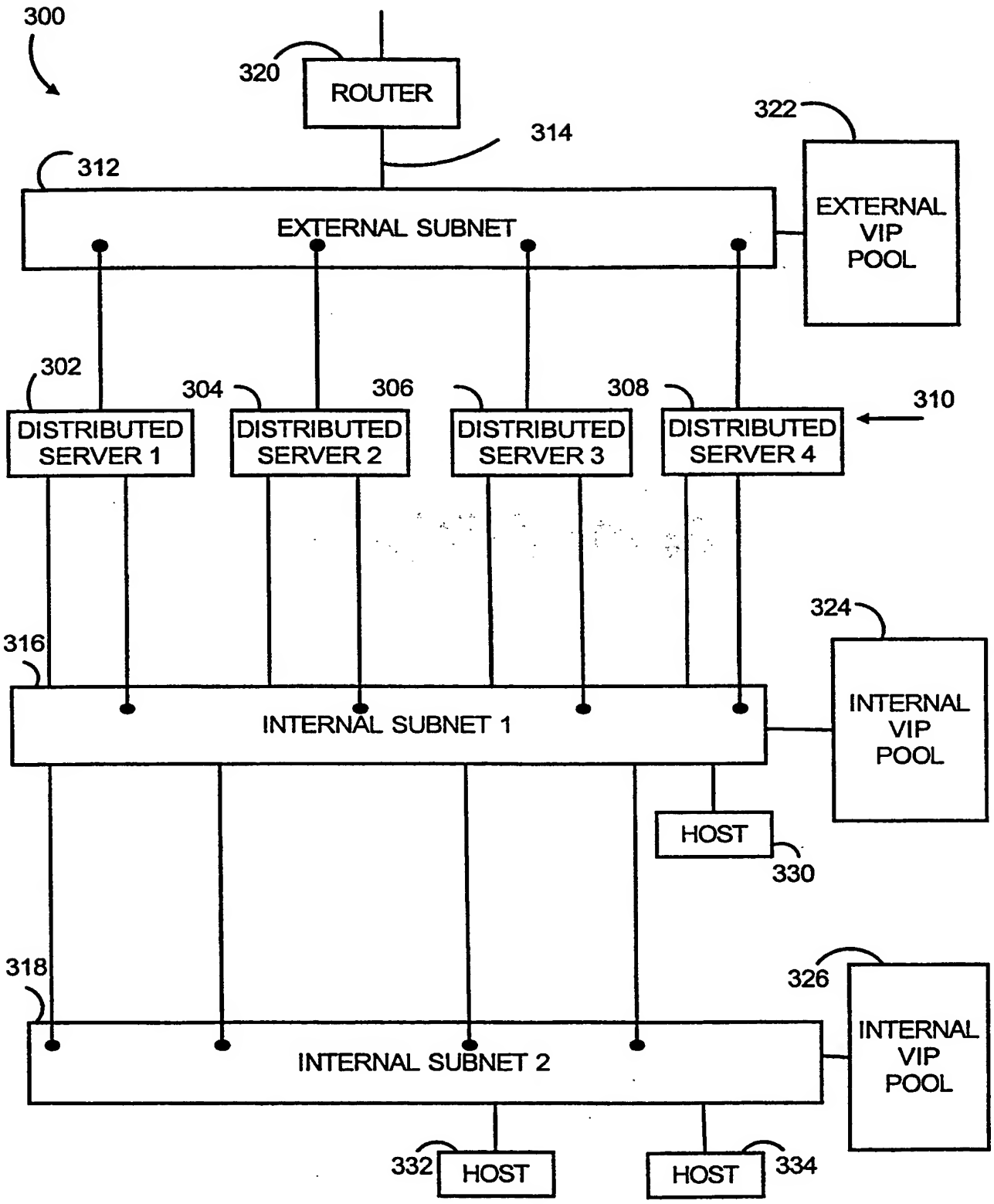

300
320
ROUTER
314
322
312
EXTERNAL SUBNET
EXTERNAL VIP POOL
302
304
306
308
310
DISTRIBUTED SERVER 1
DISTRIBUTED SERVER 2
DISTRIBUTED SERVER 3
DISTRIBUTED SERVER 4
324
316
INTERNAL SUBNET 1
INTERNAL VIP POOL
HOST
330
318
326
INTERNAL SUBNET 2
INTERNAL VIP POOL
332
HOST
HOST
334

FIG. 3

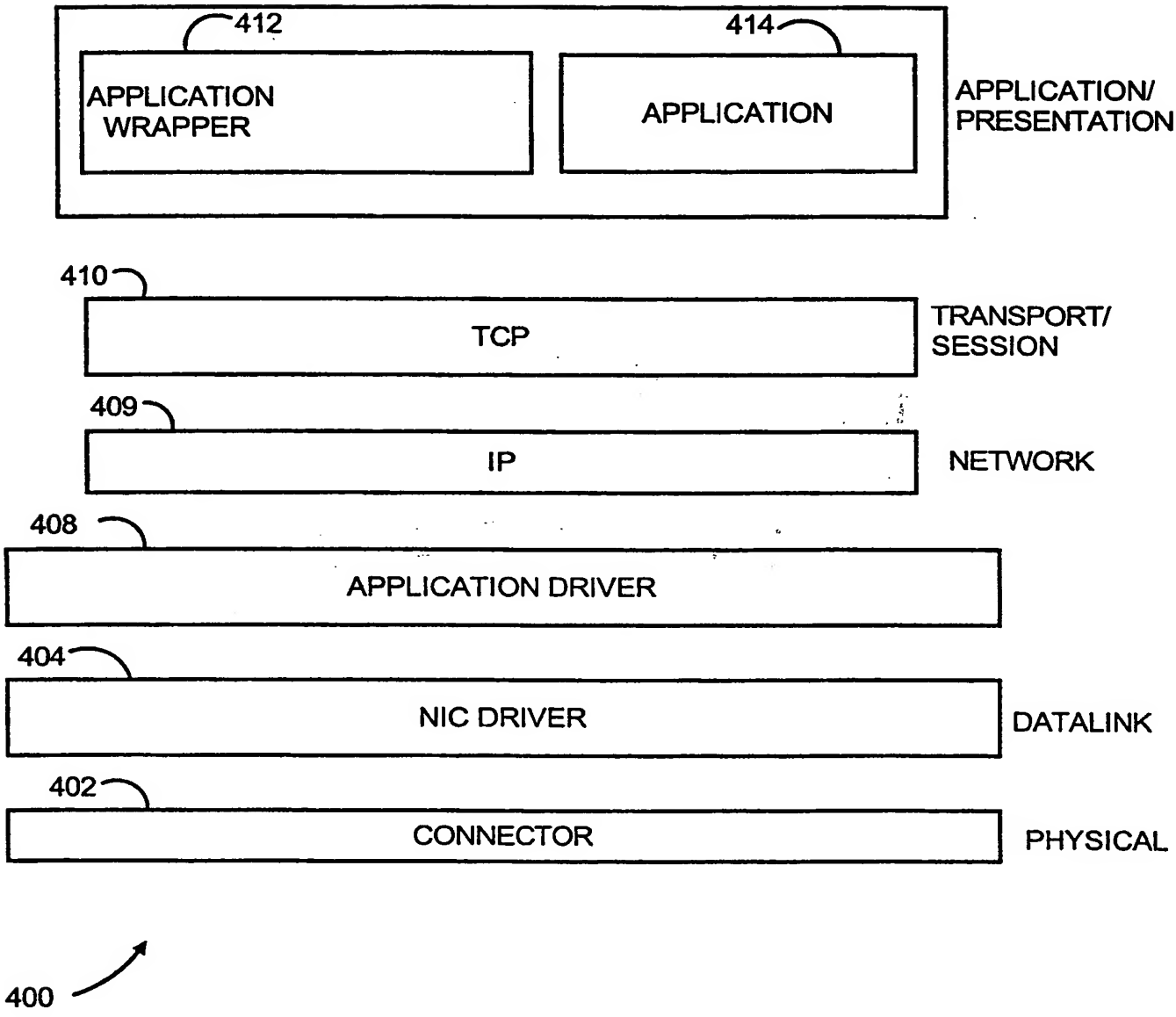

412
414
APPLICATION WRAPPER
APPLICATION
APPLICATION/ PRESENTATION
410
TCP
TRANSPORT/ SESSION
409
IP
NETWORK
408
APPLICATION DRIVER
404
NIC DRIVER
DATALINK
402
CONNECTOR
PHYSICAL
400

FIG. 4

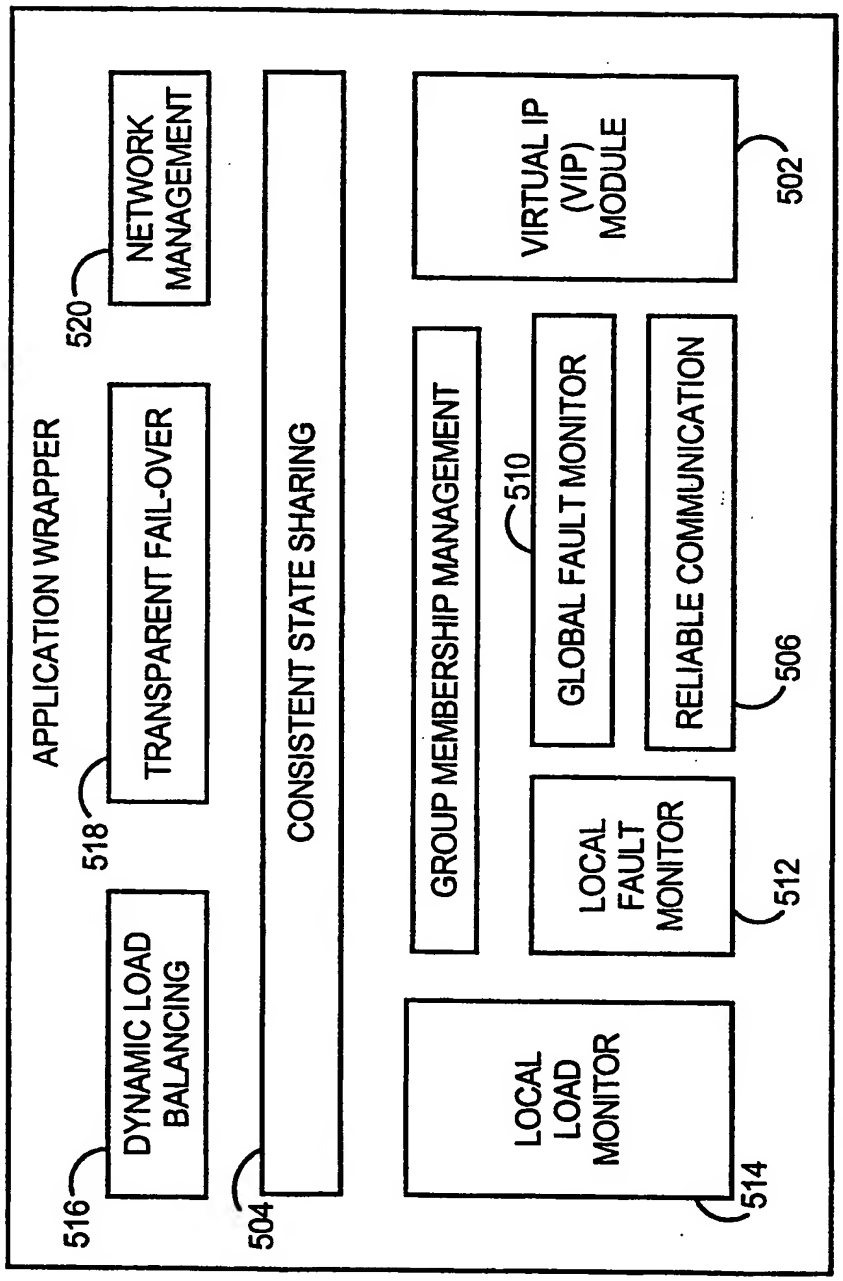


FIG. 5

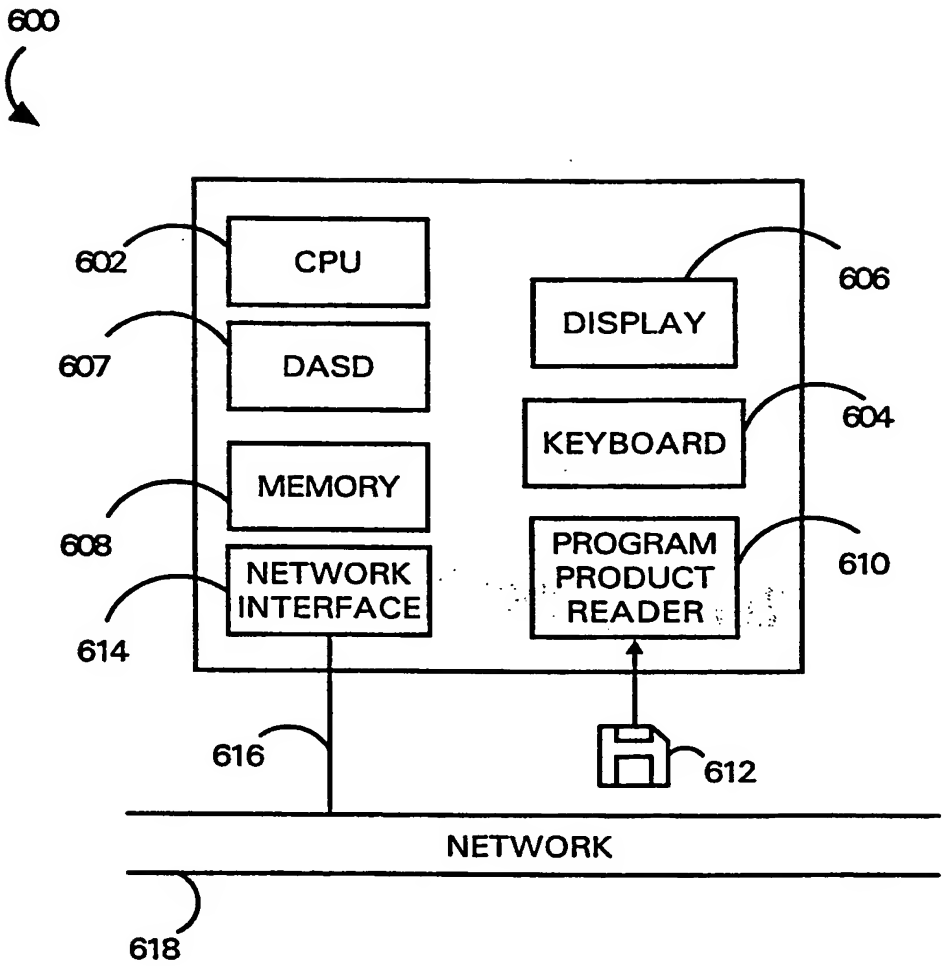
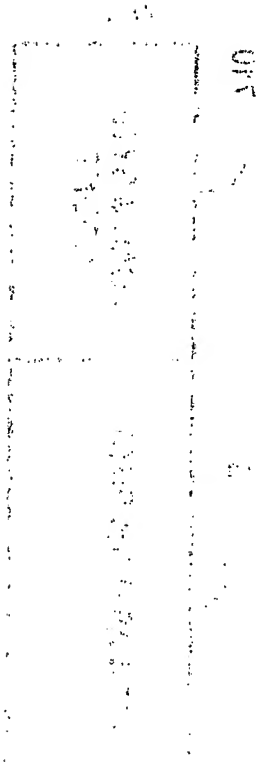

600
602
CPU
606
DISPLAY
607
DASD
604
KEYBOARD
608
MEMORY
610
PROGRAM PRODUCT READER
614
NETWORK INTERFACE
616
612
NETWORK
618

FIG. 6



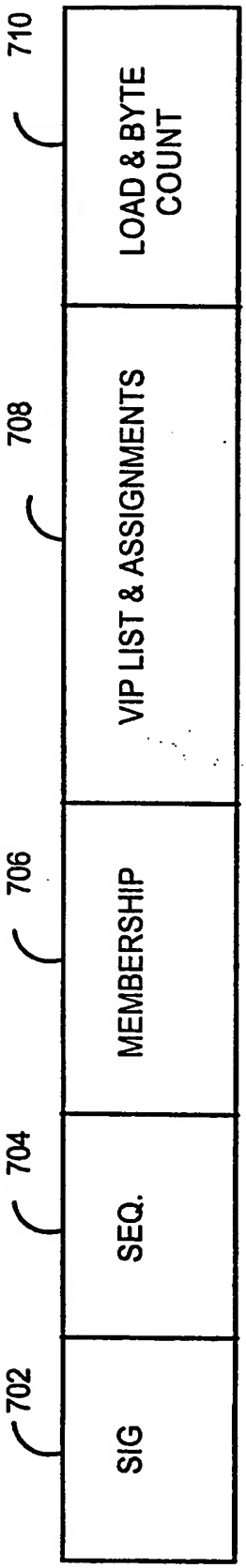
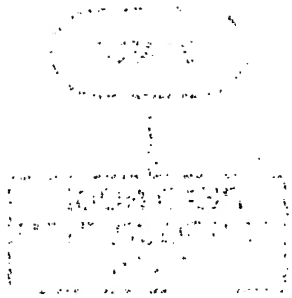

710
LOAD & BYTE COUNT
708
VIP LIST & ASSIGNMENTS
706
MEMBERSHIP
704
SEQ.
702
SIG
700

FIG. 7



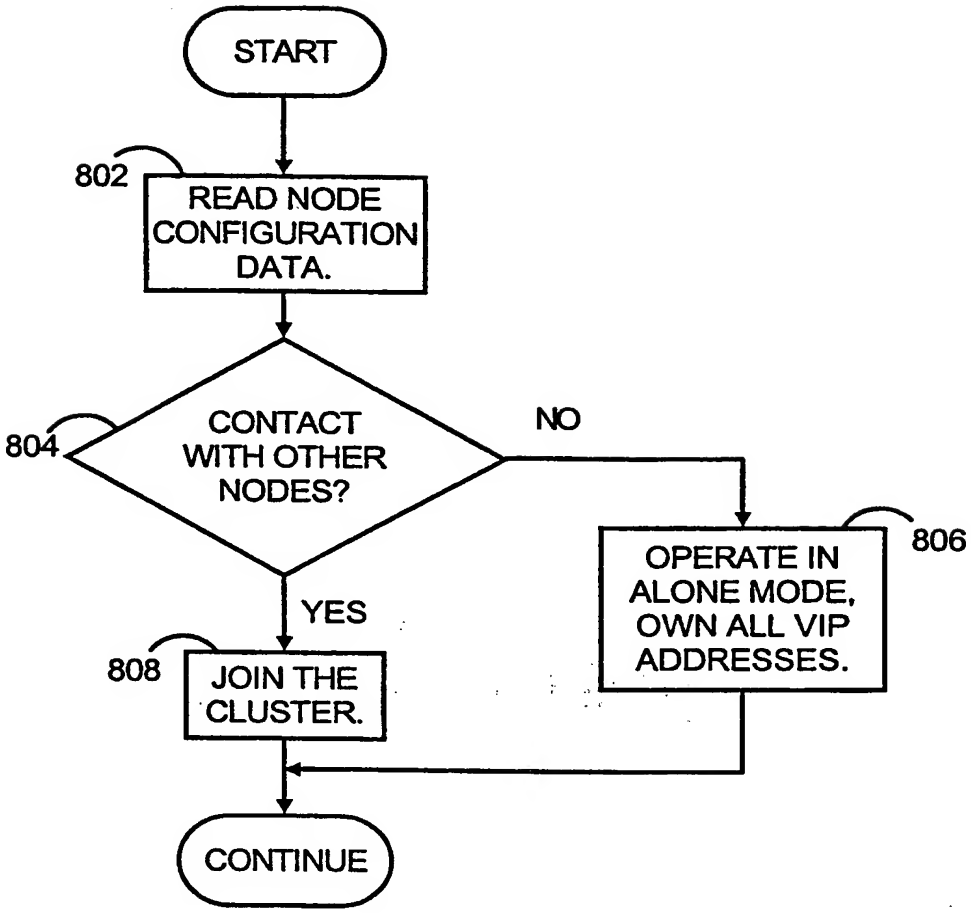


FIG. 8

This Page Blank (uspto)

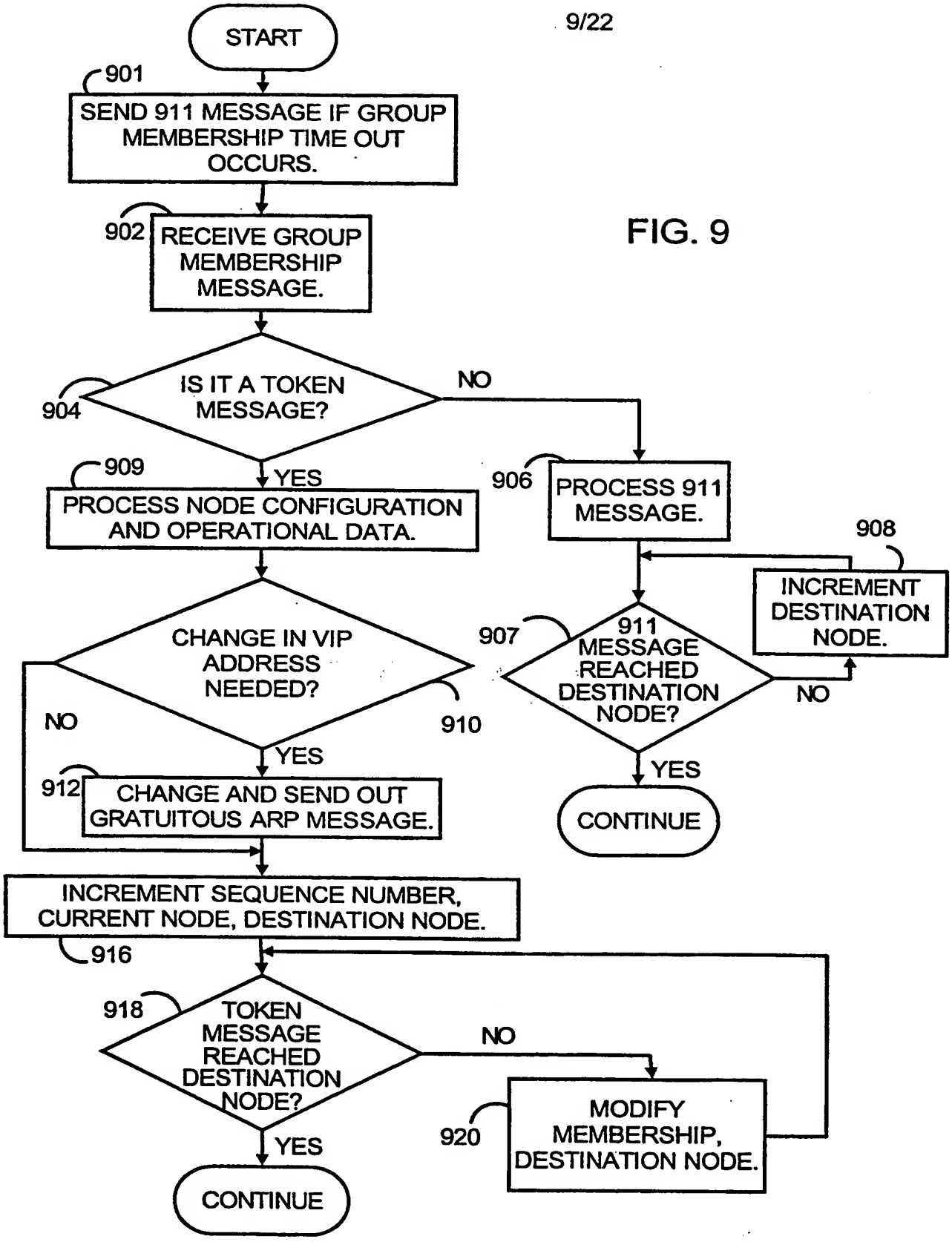

START
901
SEND 911 MESSAGE IF GROUP MEMBERSHIP TIME OUT OCCURS.
902
RECEIVE GROUP MEMBERSHIP MESSAGE.
904
IS IT A TOKEN MESSAGE?
NO
YES
909
PROCESS NODE CONFIGURATION AND OPERATIONAL DATA.
906
PROCESS 911 MESSAGE.
908
INCREMENT DESTINATION NODE.
907
911 MESSAGE REACHED DESTINATION NODE?
NO
YES
CONTINUE
CHANGE IN VIP ADDRESS NEEDED?
910
NO
YES
912
CHANGE AND SEND OUT GRATUITOUS ARP MESSAGE.
INCREMENT SEQUENCE NUMBER, CURRENT NODE, DESTINATION NODE.
916
918
TOKEN MESSAGE REACHED DESTINATION NODE?
NO
YES
CONTINUE
920
MODIFY MEMBERSHIP, DESTINATION NODE.

FIG. 9

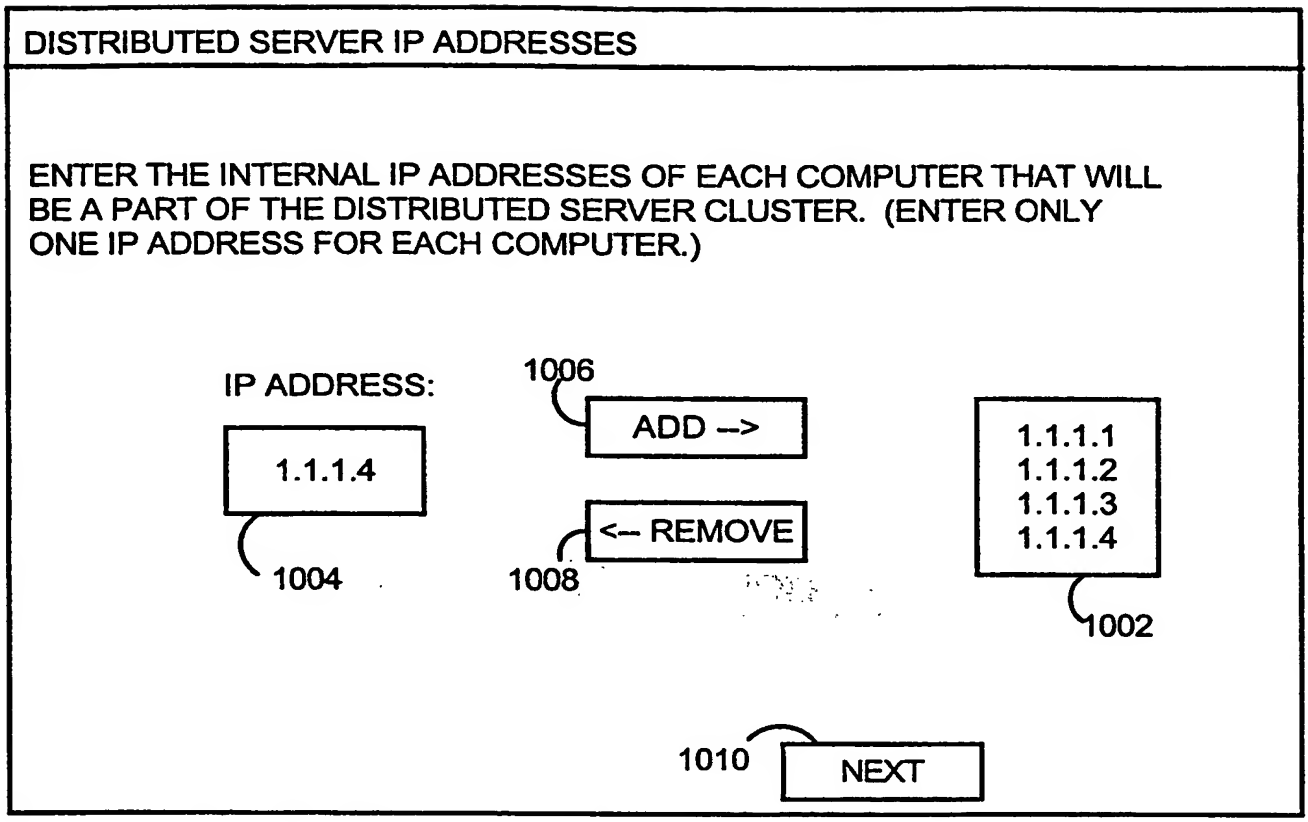

DISTRIBUTED SERVER IP ADDRESSES
ENTER THE INTERNAL IP ADDRESSES OF EACH COMPUTER THAT WILL BE A PART OF THE DISTRIBUTED SERVER CLUSTER. (ENTER ONLY ONE IP ADDRESS FOR EACH COMPUTER.)
IP ADDRESS:
1.1.1.4
1004
1006
ADD -->
<-- REMOVE
1008
1.1.1.1
1.1.1.2
1.1.1.3
1.1.1.4
1002
1010
NEXT

FIG. 10

1000

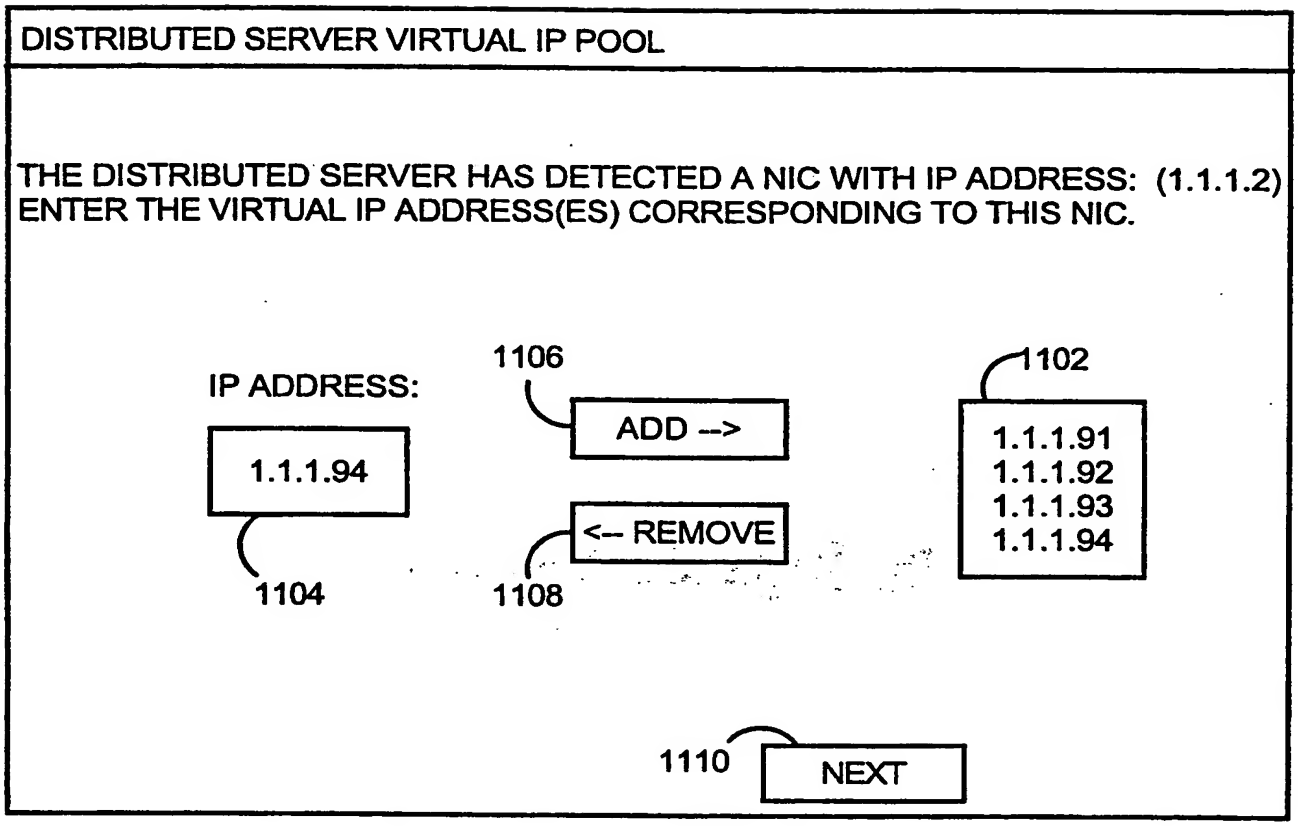


FIG. 11

1100

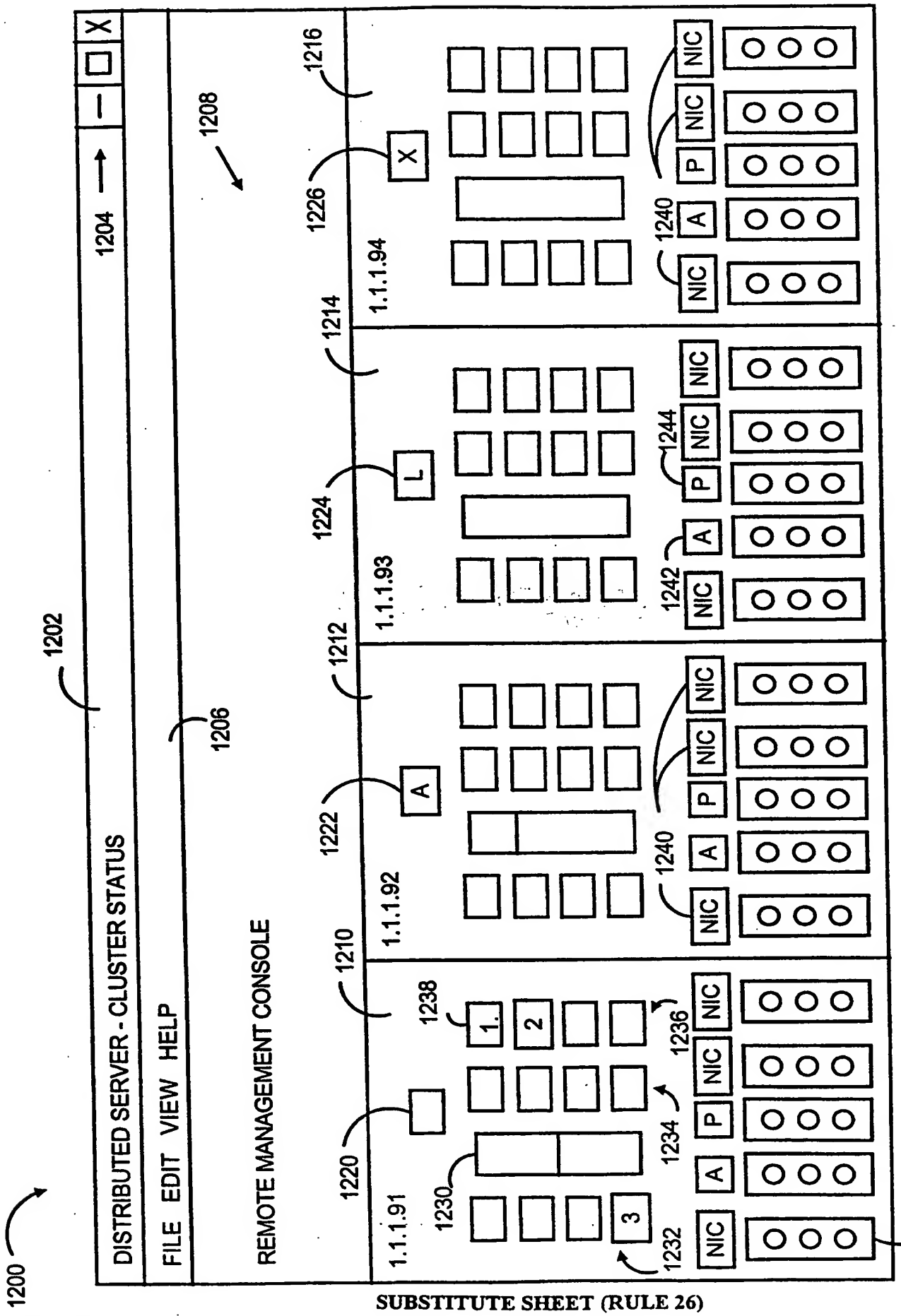


FIG. 12

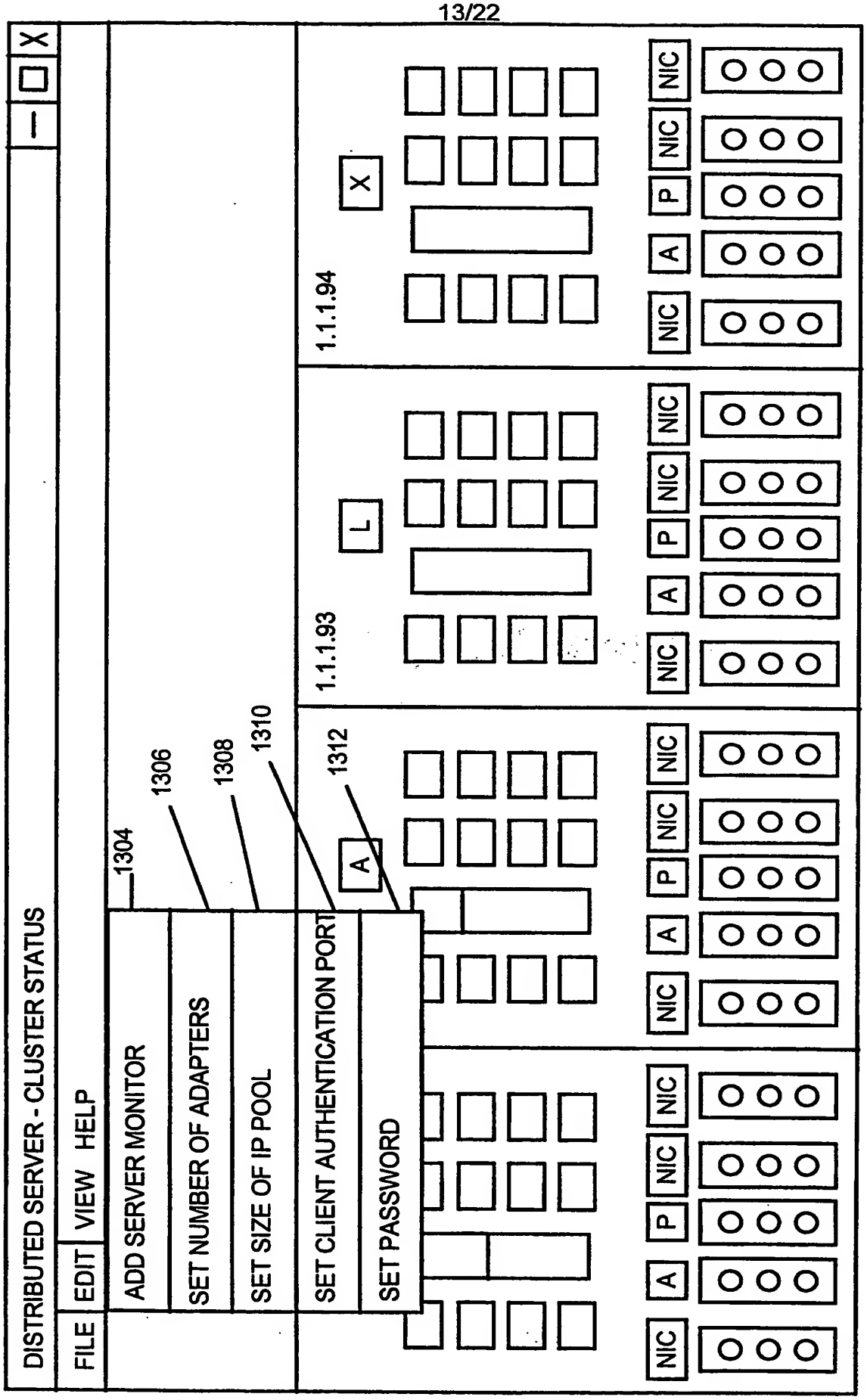


FIG. 13

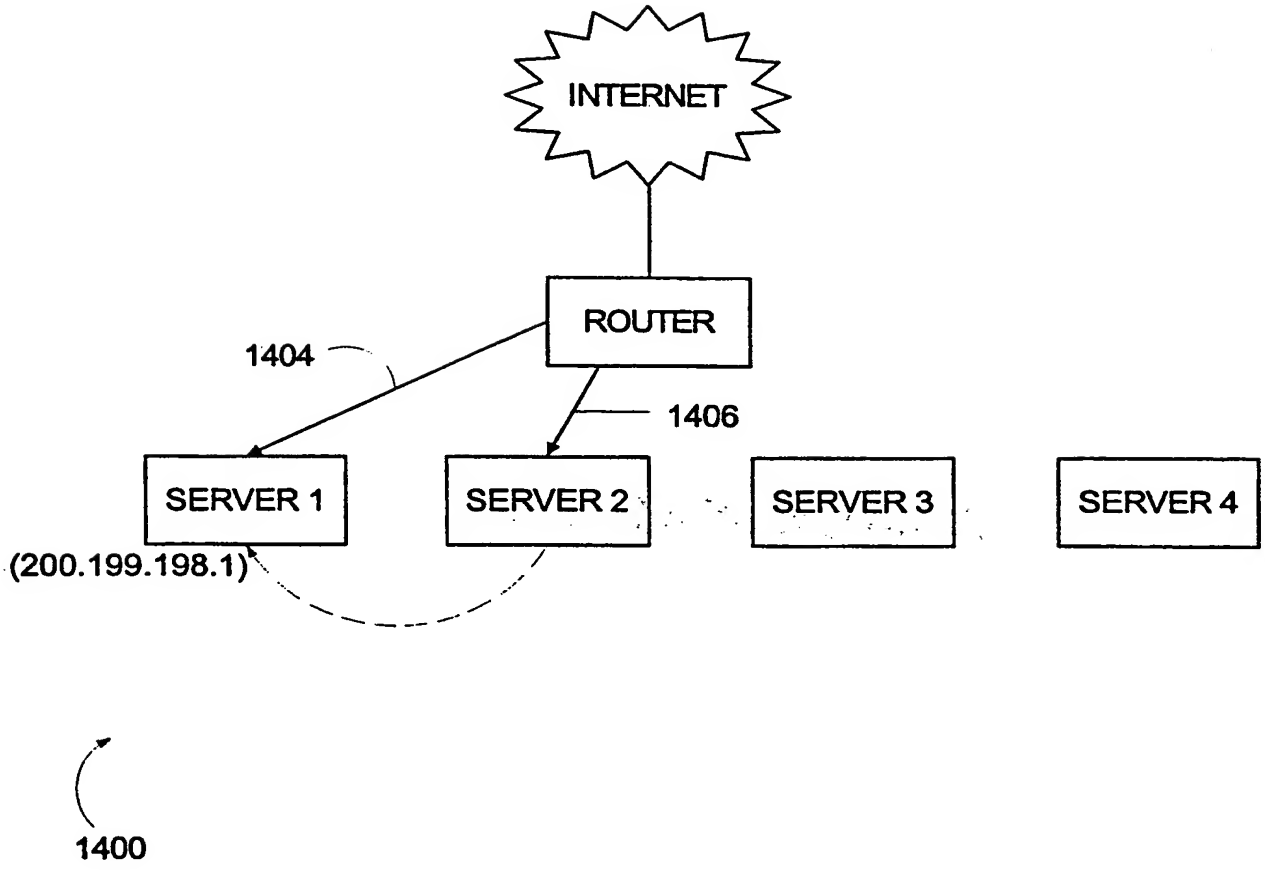

INTERNET
ROUTER
1404
1406
SERVER 1
SERVER 2
SERVER 3
SERVER 4
(200.199.198.1)
1400

FIG. 14

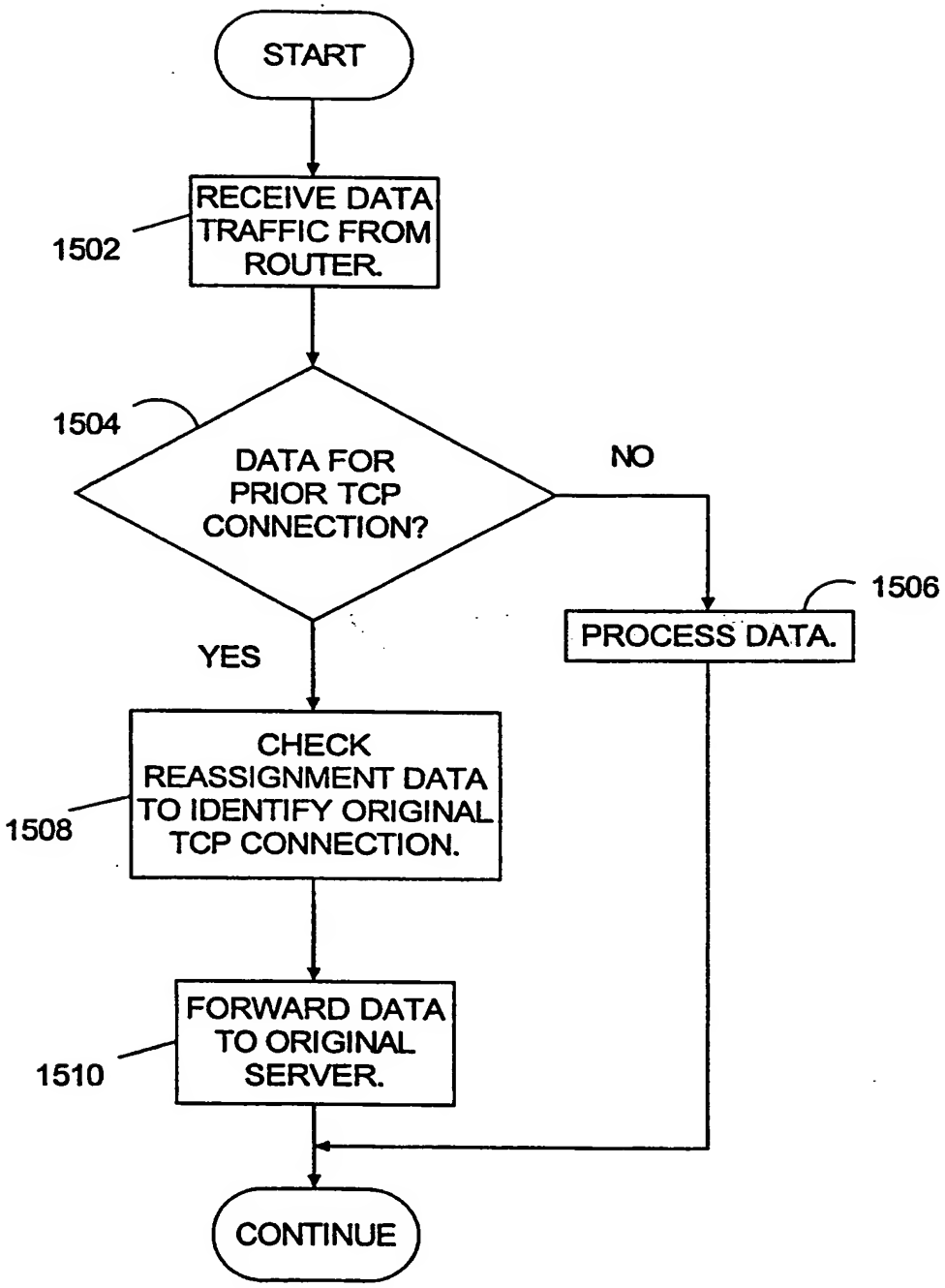


FIG. 15

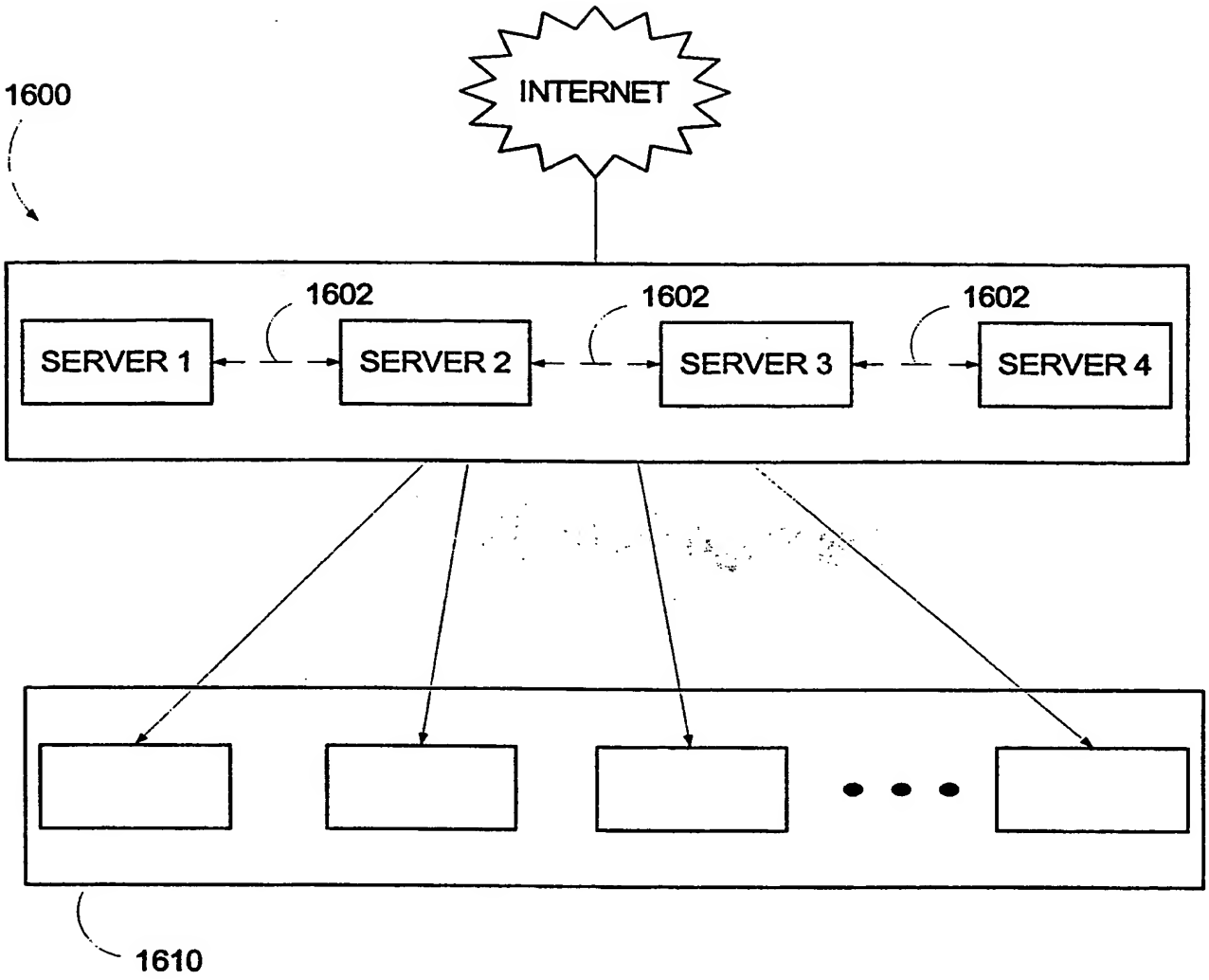

1600
INTERNET
1602
1602
1602
SERVER 1
SERVER 2
SERVER 3
SERVER 4
1610

FIG. 16

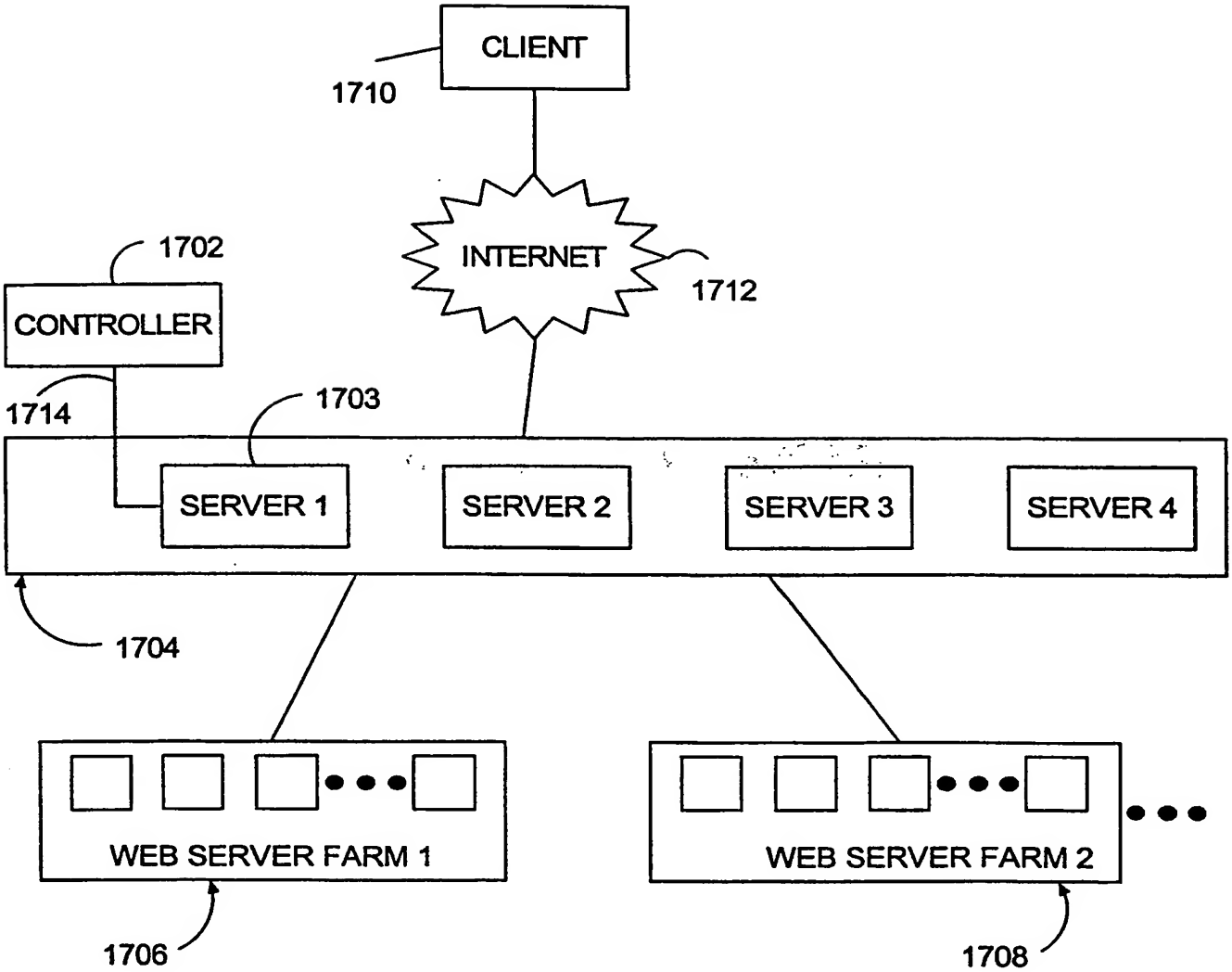
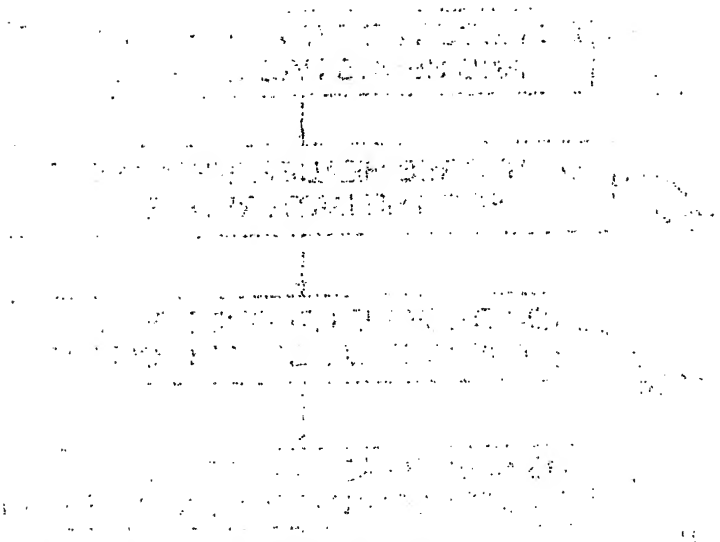

CLIENT
1710
INTERNET
1712
1702
CONTROLLER
1714
1703
SERVER 1
SERVER 2
SERVER 3
SERVER 4
1704
WEB SERVER FARM 1
WEB SERVER FARM 2
1706
1708

FIG. 17



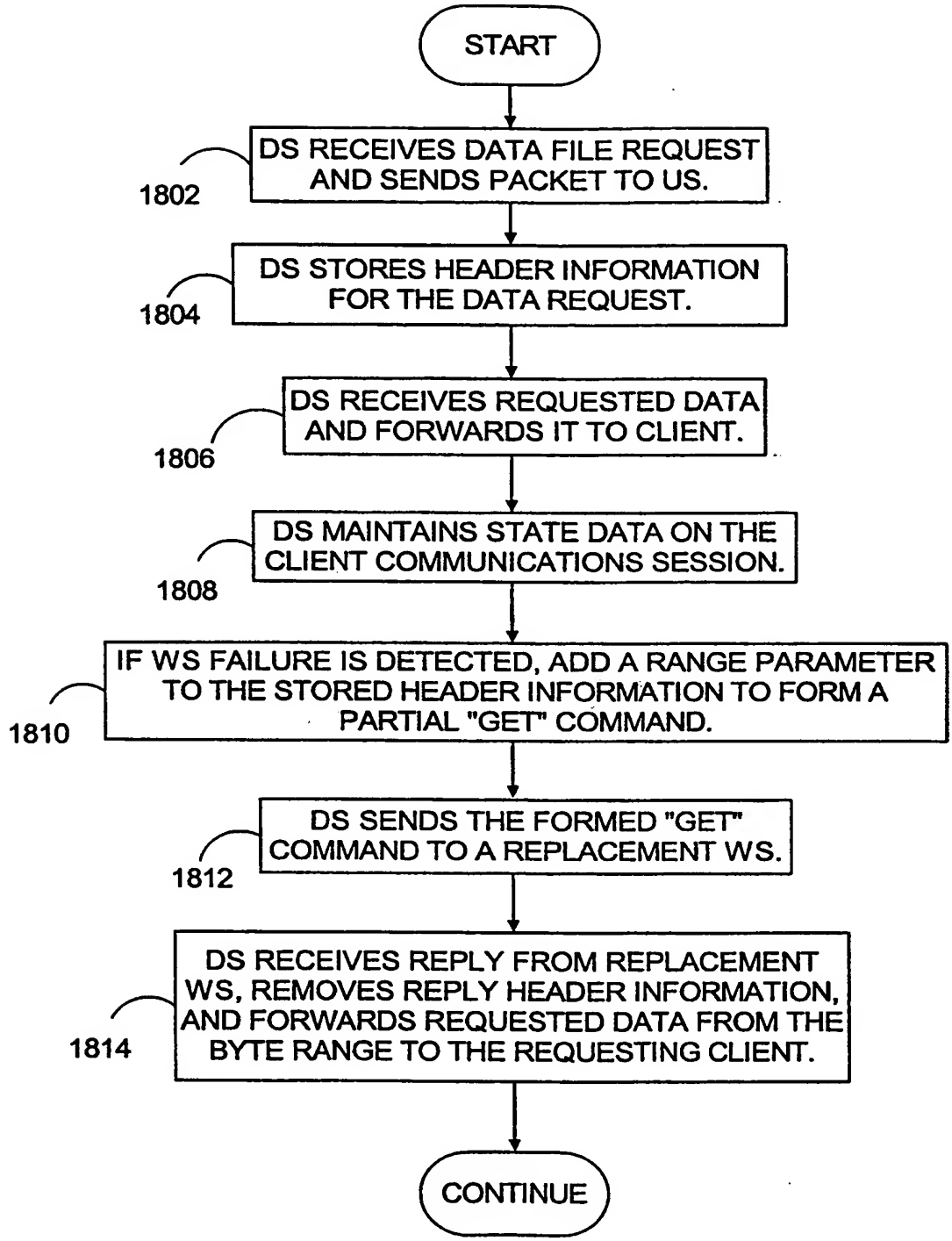
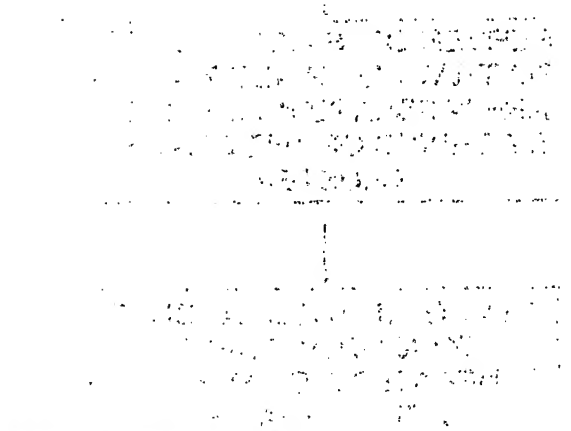


FIG. 18



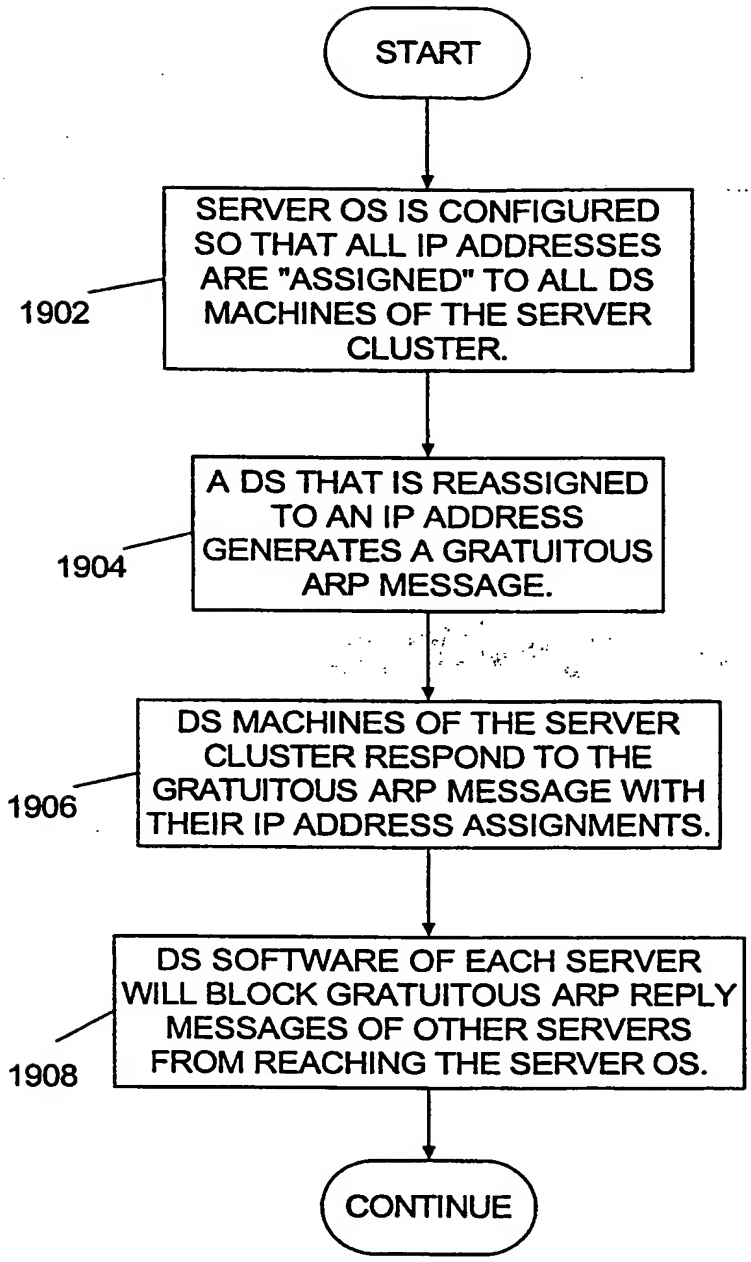


FIG. 19

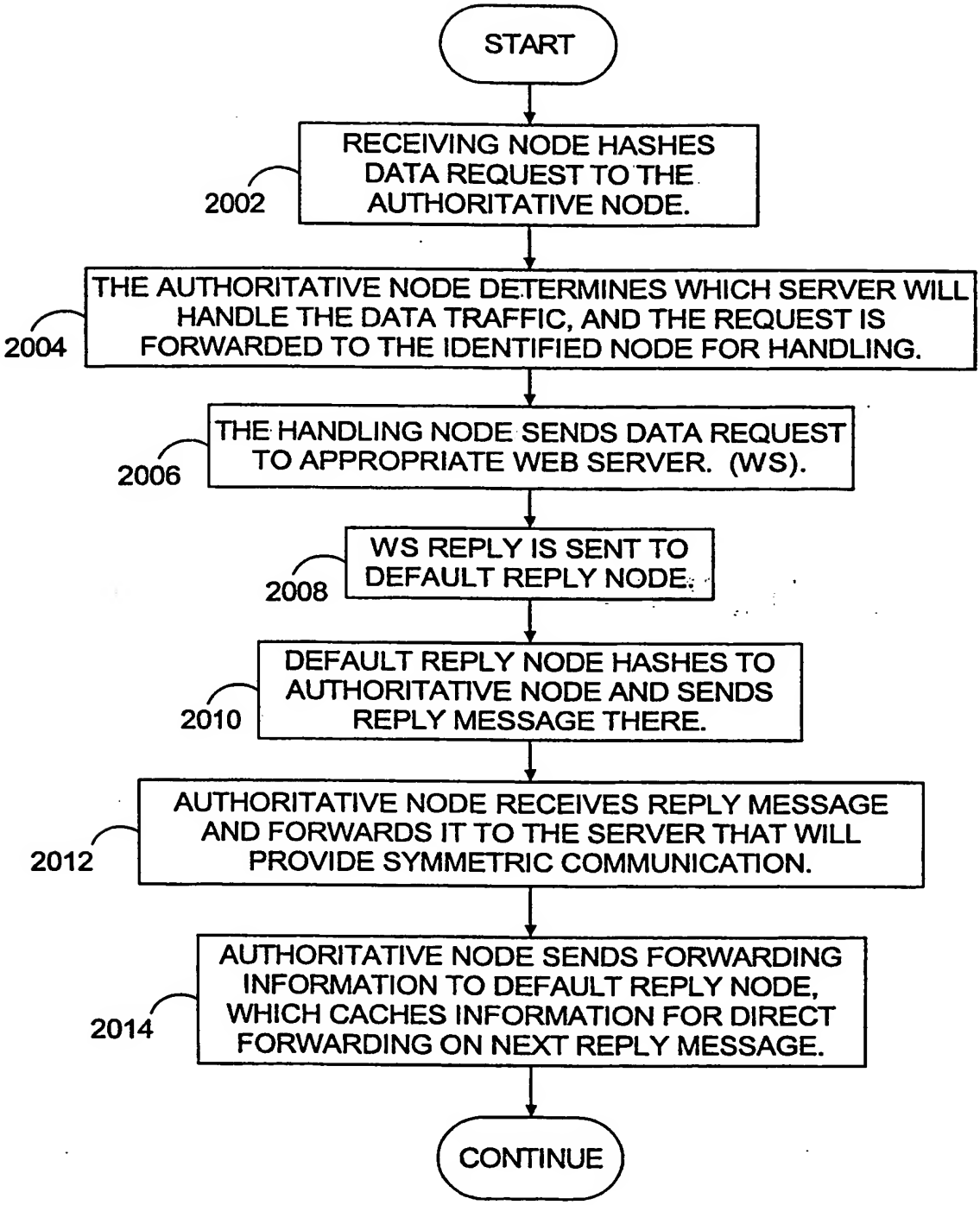


FIG. 20

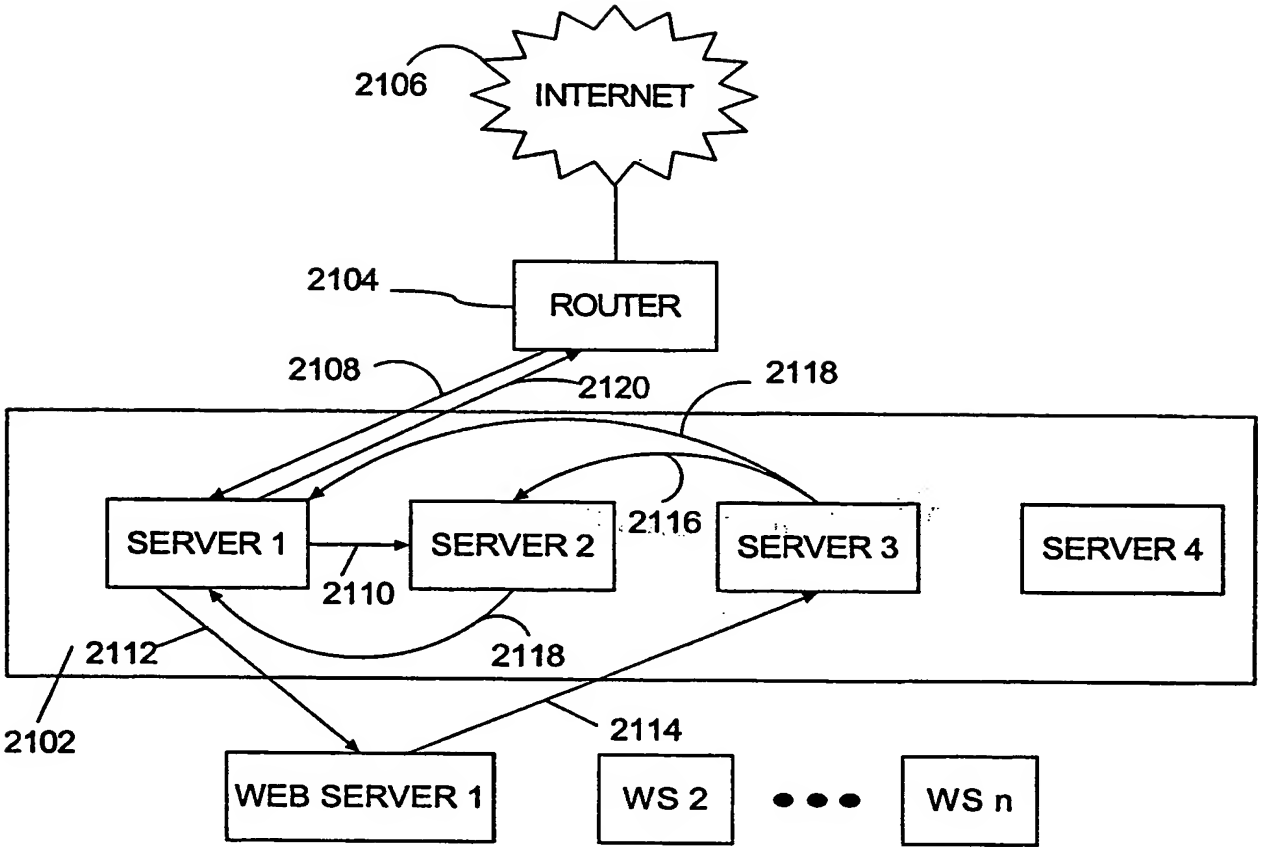

2106
INTERNET
2104
ROUTER
2108
2120
2118
SERVER 1
SERVER 2
2116
SERVER 3
SERVER 4
2110
2112
2118
2102
2114
WEB SERVER 1
WS 2
WS n

FIG. 21

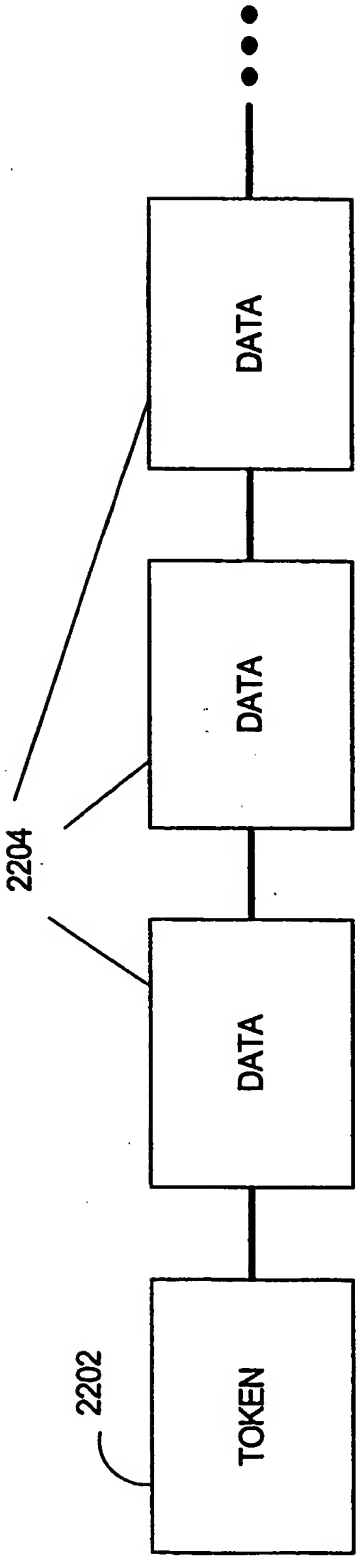


FIG. 22